心理学报 2023, Vol. 55, No. 6, 877−891
Acta Psychologica Sinica

https://doi.org/10.3724/SP.J.1041.2023.00877

奖赏预测误差对项目和联结记忆影响的分离：元记忆的作用*

龙翼婷[1] 姜英杰[1] 崔 璨[2] 岳 阳[1]

([1]东北师范大学心理学院，长春 130024)
([2]江苏省脑疾病与生物信息重点实验室，徐州医科大学生物化学与分子生物学研究中心，徐州 221004)

摘 要 选取奖赏预测误差(reward prediction error, RPE)效价和凸显性为自变量，通过3个实验考察RPE对项目和联结记忆影响的差异及其元记忆机制。被试在对图片的奖赏猜测−奖赏反馈中形成 RPE，且需要同时记忆图片(项目)以及图片−奖赏联结，最后进行记忆测试。结果表明，(1)联结记忆成绩存在RPE正效价和低凸显性优势，其信心判断准确性在RPE正效价时更高，而项目记忆成绩存在RPE负效价和高凸显性优势；(2)在编码过程中，RPE正效价和低凸显性提高了个体的瞳孔变化均值和峰值；RPE低凸显性增加了分值注视时间，缩短了图片注视时间；(3)增加RPE水平后，RPE对项目和联结记忆成绩的分离影响仍稳定存在。这些结果表明，RPE对项目和联结记忆的影响存在分离，编码阶段中，个体以 RPE 效价和凸显性为线索，通过元记忆控制对项目和联结记忆加工中的认知资源进行差异性分配；提取阶段中，RPE正效价提高了对联结记忆提取的元记忆监测水平。

关键词 奖赏预测误差，联结记忆，眼动，情景记忆，元记忆
分类号 B842

1 引言

情景记忆(episodic memory)是指个体对亲身经历过的发生在特定时间和地点的事件的记忆(Tulving, 2001)。其中的项目记忆(item memory)是对情景中单个项目的记忆，而联结记忆(associative memory)需要通过绑定加工(binding)将多个项目或特征进行整合(Murty et al., 2016)。在日常生活中，认知资源的有限性使个体难以对情景中的全部信息进行加工，为了使此刻的经验有助于将来的行动，个体通常会在编码中把认知资源分配给更有价值的信息，因此同时呈现的项目和联结记忆信息可能会因认知资源竞争被差异性加工。

有研究发现奖赏结果引起了个体对项目和联结记忆的差异性加工。Murty 等(2016)使用独裁者范式，学习阶段的每个试次中，独裁者拥有$10，可以分配给自己和被试，告知被试其中某个独裁者所分配金额将被付现，且之后还将见到这些独裁者面孔。为获得更多收益，被试在学习阶段需利用有限的认知资源，同时对面孔(项目)及其价值信息(项目−奖赏联结)进行记忆，能够体现奖赏如何影响情景中的不同记忆成分。记忆测试的结果表明，在独裁者事件中，项目记忆受奖赏影响不显著，但低奖赏结果促进了个体联结记忆的提高。

另有研究考察了奖赏预测误差(reward prediction error, RPE)对项目记忆和联结记忆的影响(Ergo et al., 2021; Rouhani et al., 2018; Rouhani et al., 2020)。RPE是奖赏结果减去奖赏预期后的差值，有效价和凸显性两个属性(Mason et al., 2019)。效价(valance)代表奖赏结果高于还是低于奖赏预期，正效价代表奖赏结果优于奖赏预期，负效价代表奖赏结果低于奖赏预期(Ergo et al., 2020)。凸显性(salience)代表

收稿日期：2022-04-14
* 吉林省自然科学基金面上项目(20230101149JC)和国家自然科学基金面上项目(32271095)资助。
通信作者：姜英杰, E-mail: jiangyj993@nenu.edu.cn

奖赏结果和奖赏预期间的偏离程度，可以用无符号的奖赏预测误差(unsigned RPE, URPE) 表示，URPE 越大代表结果与预期的偏离程度越大，个体的意外程度越高(Rouhani et al., 2018)。RPE 的效价和凸显性都会对情景记忆产生影响。

Rouhani 等(2018)在学习阶段让被试对图片进行奖赏预期并获得奖赏反馈结果，习得图片的奖赏规律(其中一类图片出现高奖赏的比率更大)，告知被试在随后的决策阶段中需要在每两张图片中进行选择并获得其奖赏；决策阶段后对被试进行了记忆测试。以反馈结果与奖赏预期之差作为 RPE，以 URPE 作为 RPE 凸显性，结果表明，RPE 对项目记忆(图片再认)影响不显著，但 URPE 显著提高了项目记忆成绩，说明项目记忆只受到了 RPE 凸显性的影响。之后 Rouhani 等(2020)将学习阶段换成被动的顺序奖赏任务，屏幕上会连续呈现若干带有不同分值的图片，分值的 10%被累计到被试收益中，由分值分布的变化产生 RPE，仍然发现了项目记忆的 RPE 高凸显性优势。此外，该研究在实验 4 中让被试对图片呈现时序进行记忆，最后的时序联结记忆成绩表现出 RPE 低凸显性优势。表明 RPE 凸显性可能引起了对项目和联结记忆的差异性加工。

但在上述研究中，只有项目记忆能影响决策收益，因此在编码阶段，被试将以 RPE 凸显性为线索，对图片编码优先投入认知资源，其次才对图片时序进行学习，这意味着联结记忆成绩可能主要取决于项目记忆对认知资源的占用情况，而非受到 RPE 的直接影响。不同的是，一些研究将联结记忆作为主要任务(外语−母语词对联结：Calderon et al., 2021；人物−物品联结：Aberg et al., 2017)，结果都发现了联结记忆中的 RPE 正效价优势，却并未能验证 RPE 凸显性的影响。Aberg 等(2017)虽然发现预期与结果偏差更小(RPE 凸显性更小)的条件下被试的联结记忆成绩更高，但实验中该条件下得到正性结果的几率也更大(RPE 效价更正)，因此难以分辨出联结记忆成绩的提高是否能归因于凸显性的影响。

综上，已有研究发现奖赏预测误差对项目和联结记忆的影响存在差异，项目记忆表现出高凸显性优势，联结记忆表现出正效价优势，但凸显性是否对联结记忆存在影响还有待考察。人类情景中通常同时涉及项目和联结成分，它们需要竞争有限的认知资源，奖赏会改变个体对这些成分的认知资源投入，因而项目和联结记忆受奖赏的影响表现出差异性(Murty et al., 2016)。但在探究 RPE 对情景记忆的影响时，先前研究对项目和联结记忆分开考察，实验往往只包含单一记忆任务，难以探讨 RPE 对项目和联结记忆影响的差异，更无法进一步揭示该差异产生的原因——即 RPE 对认知资源分配的影响。因此本研究参考 Murty 等(2016)的研究，以图片和图片−奖赏联结分别作为项目和联结记忆指标，考察个体如何基于 RPE 同时进行项目记忆和联结记忆加工。

记忆编码中个体对认知资源投入的调节，是元记忆控制过程的体现，为揭示奖赏对这一过程的影响，价值导向元记忆(value-directed metamemory)研究中采用眼动技术追踪被试对特定兴趣区中的刺激的注视时间(fixation duration)作为学习时间分配(study-time allocation)的指标，发现高价值刺激比低价值刺激获得了更长的学习时间，表明个体会优先对高价值项目分配认知资源(姜英杰 等, 2016)。瞳孔扩张也与认知资源投入有关，学习者编码信息的心理努力更大时(此时认知资源投入更多)，瞳孔扩张会增大，记忆效果提高(Ariel & Castel, 2014)。因此本研究采用眼动追踪技术，通过注视时间和瞳孔直径变化两个指标，考察 RPE 影响项目和联结记忆编码的元记忆控制过程，作为 RPE 效应产生的认知机制。

因此，本研究通过 3 个实验考察 RPE 的效价和凸显性对同时呈现的项目记忆和联结记忆的影响及其元记忆机制。其中实验 1 比较了不同 RPE 水平下项目和联结记忆成绩的差异，并通过信心判断考察了 RPE 对提取阶段元记忆监测的影响，基于 Rouhani 等(2020)的结果，实验 1 假设项目和联结记忆受 RPE 影响的方向相反，且 RPE 能够促进记忆提取中的信心判断准确性。在实验 1 的基础上，实验 2 通过眼动技术考察编码阶段中 RPE 如何影响元记忆控制过程，我们预期，项目编码和联结编码存在认知资源竞争，RPE 对二者的影响方向与记忆成绩的结果相符。在实验 1 和 2 中，奖赏结果只有 1、4 和 7 三个水平，RPE 正负效价分别与 7 和 1 两种结果存在较大的重叠，为减小 RPE 效价和奖赏结果影响的重叠，增加结果的可重复性，补充进行实验 3，在实验 1 的基础上增加 RPE 水平，验证记忆中的 RPE 效应，预期其结果会与前两个实验一致。

2 实验 1：奖赏预测误差对项目和联结记忆成绩的影响

实验 1 的目的是考察 RPE 效价和凸显性如何影

响项目和奖赏联结记忆。

2.1　方法

2.1.1　被试

招募被试 36 人, 剔除了其中 2 名(项目记忆成绩低于随机水平), 保留被试 34 人(女性 24 人), 年龄 19~25 岁(M = 22.11 岁, SD = 1.98 岁), 视力或矫正视力正常, 无神经系统和心理疾病病史。所有被试均自愿参加实验, 并签署被试知情同意书。每名被试获得真实现金奖励平均约为 21 元。

2.1.2　实验材料

刺激材料为 120 张室内和户外场景图片(Höeltje & Mecklinger, 2020), 整个实验过程中的刺激均呈现在白色背景上。通过 E-prime 3.0 进行实验编程。

考虑到个体对奖赏加工的差异能够调节 RPE 效价对记忆的影响(Aberg et al., 2017; Rouhani & Niv, 2019), 在学习和测试的间隔时间内, 让被试填写汉化版的奖惩敏感性量表(sensitivity to punishment and sensitivity to reward questionnaire, SPSRQ)。SPSRQ 包括惩罚敏感性(SP)和奖励敏感性(SR)两个分量表, 要求被试做出"是"或"否"的回答。汉化版 SPSRQ 由郭永香等(2011)进行修订, 共 31 道题目, 其中 SP 有 19 个项目, SR 有 12 个项目。汉化版 SPSRQ 中 SP 和 SR 在本研究中的内部一致性信度分别为 0.85 和 0.70。

2.1.3　实验设计

采用被试内设计, 自变量为 RPE 效价和凸显性, 因变量为项目和联结记忆成绩。

实验 1 中奖赏结果包含 1、4 和 7 三种分值, 它们在高、低价值图片中的比例分别是 2 : 3 : 5 和 5 : 3 : 2。由奖赏结果和被试的给出的猜测价值之差产生 RPE, 共有−6、−3、0、3 和 6 五种分值, 效价为其正负符号, 凸显性为 URPE。

2.1.4　实验流程

包含价值学习和记忆测试两个阶段, 总体流程如图 1。参考 Rouhani 等(2018)的实验范式, 价值学习阶段以室内和户外风景图片为学习材料, 其中一类图片出现高分值结果的概率更大。在学习阶段, 让被试对图片进行奖赏预测并获得奖赏结果的反馈, 通过这样的强化学习探索图片分值规律, 所有分值的 30%将累计入总分值中; 告知被试之后有在每两张图片中进行决策并获得其价值的机会, 以此产生奖赏动机对记忆的影响, 但实际的程序并不包括决策阶段, 学习阶段后直接对被试进行项目和奖赏联结记忆测试, 并给出对其回忆结果的信心判断(judgements of confidence, JOCs)。

学习阶段。在 400~600 ms 的注视点后, 屏幕上呈现一张室内或户外场景图片 3 秒, 这 3 秒之内不需按键。图片消失后, 进入价值猜测, 电脑询问被试"你猜测这张图片分值为多少? "并给出可选项为 1、4 或 7 分(分别对应左、下、右方向键), 要求在 3 秒内按键作答。随后询问被试对于猜测的信心判断值, 并按 Z、X、C 或 V 键进行反应(分别对应"猜的"、"有点确定"、"相当确定"或"完全确定"), 要求在 3 秒内按键反应。按键之后, 屏幕上再次呈现图片及其价值结果(如"+7"), 呈现 3 秒。正式实验包括 60 个试次, 在此之前, 被试需要完成 6 个练习试次, 以确保完全理解指导语。

学习阶段之后, 要求被试填写汉化版的 SPSRQ, 然后进行连续减 3 的干扰任务, 以填充学习和测试间的 5 分钟间隔。

最后进入未被提前告知的测试阶段。首先呈现一个 400~600 ms 的注视点, 随后呈现一张场景图片, 图片下方出现新旧两个选项, 被试需要在 5 秒内按左右方向键进行反应, 并在之后的 3 秒内按 Z、X、C 或 V 键做出信心判断(猜的、有点确定、相当确定或完全确定)。然后进入价值回忆, 屏幕上再次呈现该场景图片, 图片下方出现 1、4 和 7 三个价值选项, 被试需要在 5 秒内按方向键作答, 并在之后

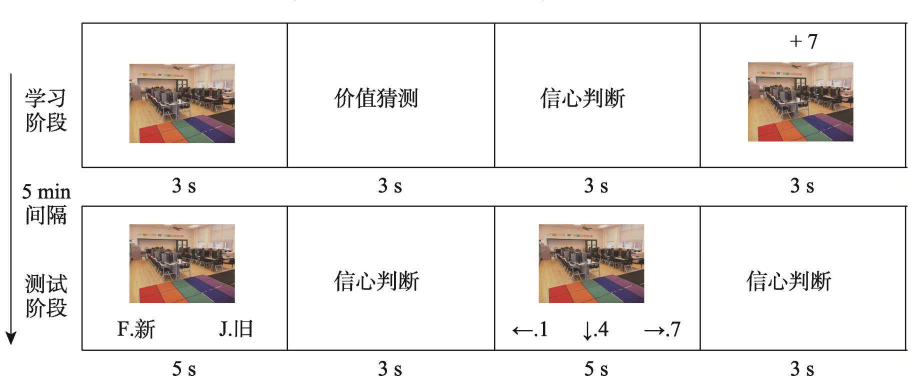


图 1　实验 1 流程图

的 3 秒内给出信心值(猜的、有点确定、相当确定或完全确定)。回忆阶段的正式实验包含 60 张旧图像和 60 张新图像, 在此之前, 被试需要完成 12 个试次的练习。

2.1.5 数据整理与分析方法

项目记忆成绩为对旧图片的再认结果(正确计分为 1, 错误计分为 0), 联结记忆成绩为项目再认正确试次中被试对价值反馈的回忆结果(正确计分为 1, 错误计分为 0), JOCs 中猜的、有点确定、相当确定、完全确定分别记分为 1、2、3、4。对于 SPSRQ 分数, 首先将 SR 分数和 SP 分数转换为 z 值, 再将二者相减得到奖惩敏感性差值分数。

参考前人的分析方法(Rouhani et al., 2018), 通过 R 软件的广义混合效应线性模型对行为数据进行分析, 始终将被试作为随机截距项放入所有模型分析。以往多数研究只将 RPE 作为预测因子, 忽略了奖赏结果本身的影响(Ergo et al., 2020; Jang et al., 2019; Rouhani et al., 2018), 但奖赏结果和 RPE 两种奖赏成分都可能影响情景记忆, 且 RPE 正效价往往对应着高奖赏结果, 两种效应存在混淆, 因此本研究将同时对这两种奖赏成分进行分析, 在考察 RPE 和 URPE 的影响时, 始终将奖赏结果作为另一预测因子放入模型。若因变量为分类变量(即项目记忆和联结记忆准确性), 使用广义混合效应线性模型(lme4 包中的 glmer 函数)进行分析; 若因变量为连续变量(即项目记忆和联结记忆 JOCs), 使用混合效应线性模型(nlme 包中的 lme 函数)进行分析。当对多个因子的影响进行分析时, 报告拟合程度最好(即 AIC 值最低)的公式中的效应, 如某些情况下, 考虑进因子间显著的交互作用后, AIC 值更低, 则同时报告因子的主效应及其交互作用。在线性模型分析结果中, RPE 的主效应显著则表明存在 RPE 效价效应, URPE 的主效应显著则表明存在凸显性效应。对于交互作用, 通过将主效应项 b 值乘积的符号与交互作用项 b 值的符号进行对比, 符号相反时的交互作用被描述为负向交互作用, 此时一个变量的增长会削弱另一个变量的效应量; 符号相同时则被描述为正向交互作用, 此时一个变量的增长会增大另一个变量的效应量。

2.2 结果分析

2.2.1 奖赏预测误差对情景记忆成绩的影响

删去学习阶段中被试未做出价值猜测按键的试次后, 不同 RPE 条件下, 项目和联结记忆击中率的均值和标准差如表 1 所示, RPE 和奖赏结果对记忆成绩的影响如图 2 所示。

对记忆成绩进行分析(表 2)发现, 奖赏结果(p = 0.610)和 RPE (p = 0.217)对项目记忆的影响不显著, URPE (p = 0.070)的影响也只达到了边缘显著水平, 表明项目记忆只有随凸显性增大而提高的趋势。对于联结记忆, 奖赏结果(p = 0.002)和 RPE (p < 0.001)主效应均显著, 表明高奖赏结果和 RPE 正效价都会促进联结记忆成绩的提高, 且两者存在显著的负向交互作用(p = 0.001), 表明随着奖赏结果的增大, RPE 效价的效应会减弱; URPE 的影响也显著(p <

表 1 实验 1 项目和联结记忆击中率均值和标准差(n = 34)

记忆类型	RPE				
	−6	−3	0	3	6
项目记忆	0.73 (0.21)	0.73 (0.16)	0.71 (0.13)	0.75 (0.17)	0.79 (0.25)
联结记忆	0.32 (0.32)	0.43 (0.17)	0.62 (0.13)	0.43 (0.18)	0.42 (0.36)

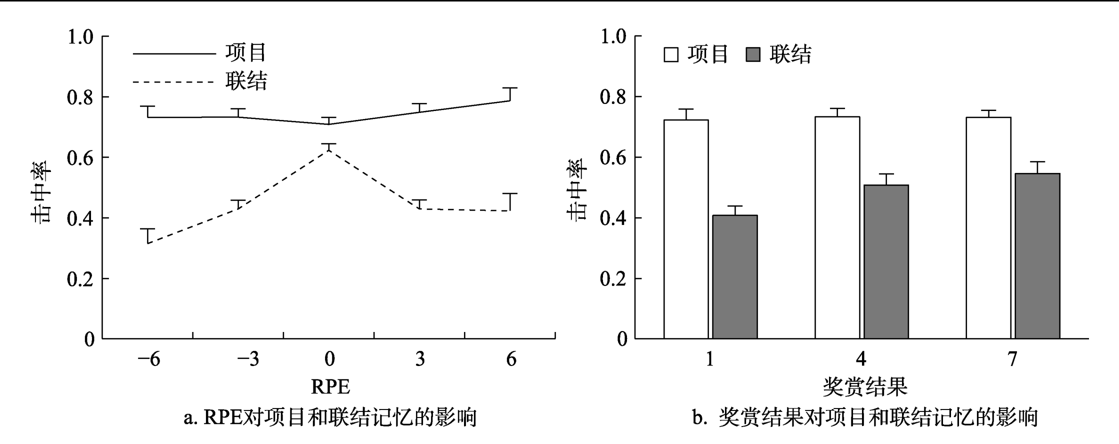


图 2 实验 1 奖赏(a 为 RPE, b 为奖赏结果)对项目和联结记忆击中率的影响(误差线为标准误)

表 2　实验 1 项目和联结记忆击中率广义混合线性模型分析结果

变量	项目记忆击中率				联结记忆击中率			
	b	*z*	*p*	95% CI	*b*	*z*	*p*	95% CI
intercept	1.11	7.18	<0.001	[0.34, 0.68]	−0.17	−1.21	0.227	[−0.46, 0.11]
RPE	0.03	1.23	0.217	[−0.02, 0.07]	0.14	3.54	<0.001	[0.06, 0.22]
奖赏结果	−0.02	−0.51	0.610	[−0.07, 0.04]	0.09	3.15	0.002	[0.04, 0.15]
交互作用					−0.04	−5.03	0.001	[−0.06, −0.02]
intercept	0.89	0.64	<0.001	[0.60, 1.17]	0.12	0.96	0.335	[−0.12, 0.36]
URPE	0.05	1.81	0.070	[−0.00, 0.09]	−0.19	−7.13	<0.001	[−0.24, −0.13]
奖赏结果	0.01	0.64	0.525	[−0.03, 0.05]	0.08	3.54	<0.001	[0.03, 0.12]

0.001), 联结记忆成绩在低凸显性时更高。

对再认正确试次的反应时进行分析(表 3), 结果只发现了奖赏结果的增大对联结记忆提取速度的显著促进作用(p = 0.002)。

将正负 RPE 效价间的项目记忆和联结记忆成绩分别作差, 再将该差值与奖惩敏感性差值进行相关分析。结果发现, 奖惩敏感性差值与项目记忆成绩差值的相关并不显著(r = −0.25, p = 0.147), 但与联结记忆成绩差值的相关显著(r = 0.39, p = 0.022), 对奖赏更敏感同时对惩罚更不敏感的被试, 在 RPE 正效价时的联结记忆成绩会更高于负效价时。

2.2.2　奖赏预测误差对信心判断的影响

删去一名未理解信心判断指导语的被试(全部选择了完全确定选项), 并删去被试在学习阶段没有做出价值预期以及在测试阶段未做出信心判断的试次后, 在不同 RPE 条件下, 不同记忆结果下中项目记忆和联结记忆 JOCs 的均值和标准差如表 4 所示, 错误和正确回忆试次中 RPE 对 JOCs 的影响如图 3 所示。

对于项目记忆 JOCs, 首先, 项目记忆结果的影响显著(b = 0.76, t = 17.48, p < 0.001, 95% CI = [0.68, 0.85]), 再认正确时的信心值比再认错误时更高。值得注意的是, 这种记忆结果的主效应体现的是 JOCs 的准确性, 若其他变量和记忆结果有显著的交互作用, 则表明该变量能够影响 JOCs 准确性。在将记忆结果作为预测因子的基础上, 分别对 RPE 和 URPE 进行混合线性回归分析(表 5)。结果发现, 三个奖赏成分与记忆结果的交互作用都不显著(ps > 0.010)。

对于联结记忆 JOCs, 联结记忆结果的影响同

表 3　实验 1 项目和联结记忆反应时广义混合线性模型分析结果

变量	项目记忆反应时(ms)				联结记忆反应时(ms)			
	b	*t*	*p*	95% CI	*b*	*t*	*p*	95% CI
intercept	1914.56	28.84	<0.001	[1784.32, 2044, 79]	1949.60	19.03	<0.001	[1748.41, 2150, 78]
RPE	8.23	−0.98	0.327	[−7, 80, 24.24]	2.13	0.16	0.874	[−23.83, 28.99]
奖赏结果	−10.23	−1.01	0.314	[−30.67, 10.22]	−49.95	−3.18	0.002	[−80.55, −18.77]
intercept	1856.27	29.24	<0.001	[1731.749, 1980, 79]	1916.96	19.47	<0.001	[1723.65, 2110.28]
URPE	10.77	1.21	0.228	[−6.74, 28.27]	11.51	0.80	0.426	[−16.88, 39.91]
奖赏结果	−2.44	−0.32	0.746	[−17.26, 12.37]	−47.31	−3.86	0.003	[−71.37, −23.24]

表 4　项目和联结记忆 JOCs 均值和标准差(n = 33)

JOCs	RPE				
	−6	−3	0	3	6
项目记忆 JOCs					
再认正确	3.07 (0.81)	3.00 (0.93)	3.02 (0.82)	2.99 (0.80)	3.21 (0.65)
再认错误	2.57 (0.75)	2.31 (0.77)	2.45 (0.6)	2.21 (0.73)	2.53 (0.91)
联结记忆 JOCs					
回忆正确	2.35 (0.87)	2.49 (0.66)	2.75 (0.55)	2.81 (0.70)	3.14 (0.85)
回忆错误	2.64 (0.81)	2.54 (0.58)	2.38 (0.52)	2.40 (0.58)	2.45 (0.69)

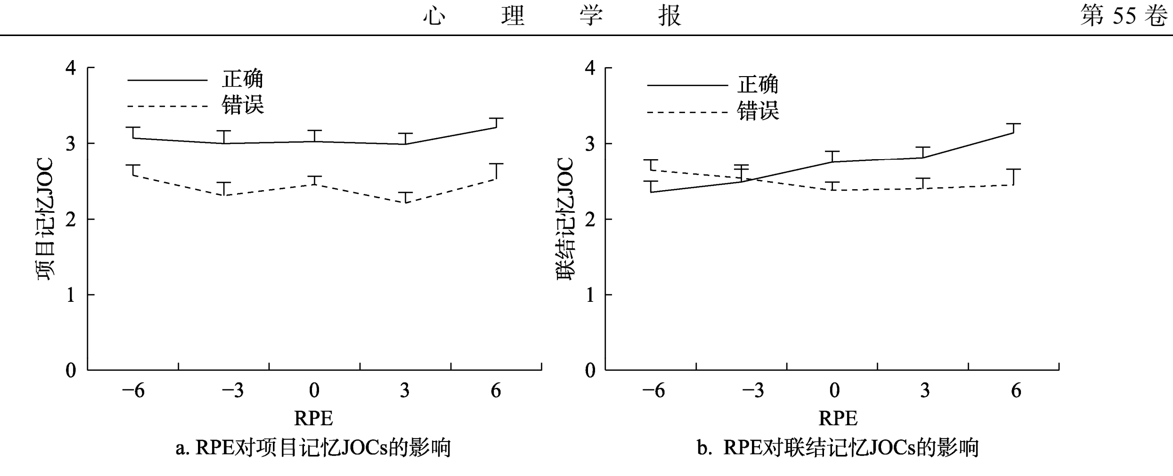


图 3　实验 1 RPE 对项目(图 a)和联结记忆(图 b)JOCs 的影响(误差线为标准误)

表 5　实验 1 JOCs 混合线性模型分析结果

变量	项目记忆 JOCs				联结记忆 JOCs			
	b	*t*	*p*	95% CI	*b*	*t*	*p*	95% CI
intercept	2.47	29.91	<0.001	[2.31, 2.63]	2.46	28.01	<0.001	[2.29, 2.64]
记忆结果	0.70	29.91	<0.001	[0.57, 0.83]	0.05	0.52	0.602	[−0.12, 0.22]
RPE×记忆结果	−0.01	−0.76	0.448	[−0.03, 0.01]	0.03	2.19	0.029	[0.00, 0.06]
奖赏结果×记忆结果	0.02	1.20	0.230	[−0.01, 0.04]	0.05	2.79	0.005	[0.01, 0.08]
Intercept	2.47	29.96	<0.001	[2.31, 2.63]	2.46	27.89	<0.001	[2.29, 2.64]
记忆结果	0.71	11.34	<0.001	[0.00, 0.09]	−0.05	−0.58	0.564	[−0.21, 0.11]
URPE×记忆结果	0.01	0.62	0.534	[−0.01, 0.03]	0.00	−0.23	0.817	[−0.04, 0.03]
奖赏结果×记忆结果	0.01	0.96	0.335	[−0.01, 0.03]	0.07	5.36	<0.001	[0.05, 0.10]

样显著(b = 0.26, t = 5.47, p < 0.001, 95% CI = [0.16, 0.35]), 当联结回忆正确时, 被试的信心判断值更高。在将回忆结果作为预测因子的基础上, 再分别对 RPE 和 URPE 的影响进行混合线性回归分析(表 5)。结果发现, 奖赏结果(p = 0.005)和 RPE (p = 0.029)分别与记忆结果的正向交互作用显著, 表明奖赏结果和 RPE 效价对联结记忆 JOCs 准确性存在显著影响, 奖赏结果升高和 RPE 正效价都促进了联结 JOCs 准确性的提高。但 URPE 与记忆结果交互作用不显著(p = 0.817)。

2.3　讨论

实验 1 中, 不同奖赏条件下的项目记忆成绩相似, 只存在项目记忆随 RPE 凸显性升高而增大的趋势。这可能是由于预期阶段时长为 3s, 导致被试在预期阶段就已对图片进行了较深编码, 而在反馈阶段更多地编码联结信息, 从而削弱了反馈阶段产生的 RPE 对项目记忆编码的影响。但前人使用相似程序(Rouhani & Niv, 2021), 却仍发现了 URPE 对项目记忆的显著影响, 究其原因, 前人研究中分值为 1 至 100 间的连续整数列, 而本研究只包含 1、4 和 7 三个分值选项, 奖赏信息加工对认知资源的占用更小, 有利于个体更充分对图片进行加工。

不同的是, 考虑了奖赏结果增大对联结记忆成绩的促进效应后, 我们仍发现了联结记忆中明显的 RPE 正效价和低凸显性优势。此外, 个体特质可能会通过影响奖赏的动机作用而改变记忆成绩(Aberg et al., 2017; Rouhani & Niv, 2019), 奖惩敏感性差值的分析结果表明, 相比于受奖赏动机的激励更弱的人, 受奖赏动机的激励更强的个体在面对 RPE 正效价时, 联结记忆成绩会更高。

反应时反映了记忆提取的流畅性, JOCs 反映了记忆提取中的元记忆监测准确性。奖赏对联结记忆的提取过程有重要影响, 一方面, 奖赏结果的增大提高了联结记忆提取的流畅性, 另一方面, 联结记忆提取中的元记忆监测在高奖赏结果和 RPE 正效价时更准确, 且 RPE 正效价时准确性也更高。

3　实验 2：奖赏预测误差对项目和联结编码中眼动的影响

在实验 1 的基础上, 实验 2 采用眼动技术, 考察 RPE 效价和凸显性对项目和联结记忆编码过程中元记忆控制的影响。实验 1 过长的预期阶段导致

了项目记忆成绩的阴性结果, 因此实验 2 中将学习阶段的预期阶段时间由 3 s 变为 2 s, 相应地将反馈阶段时间由 3 s 延长到 4 s, 以增强反馈阶段的奖赏信息对记忆的影响。

3.1 方法

3.1.1 被试

考虑到眼动分析中会剔除部分记录失误的试次, 实验 2 将学习阶段的试次数由 60 增加到 100, 相应减小了被试量。共招募被试 23 人, 其中女性 17 人, 年龄为 17~23 岁(M = 19.79 岁, SD = 1.89 岁), 视力或矫正视力正常, 无神经系统和心理疾病病史。所有被试均自愿参加实验, 并签署被试知情同意书。每名被试获得真实现金奖励平均约为 28 元。

3.1.2 实验材料与仪器

刺激材料和量表与实验 1 中相同。采用 Eyelink1000Plus 塔式眼动仪, 采样率为 1000 Hz。屏幕刷新率为 60 Hz, 分辨率为 1920×1080 像素, 被试眼睛与屏幕间距离为 76 cm, 所有刺激都呈现在深灰色背景上, 图片大小为 600×400 像素, 记录右眼的眼动轨迹。

3.1.3 实验流程

实验流程(图 4)和指导语与实验 1 相似。不同的是, 调整学习阶段各屏的呈现时间, 预期阶段图片仅呈现 2 秒, 而反馈阶段图片呈现时长改为 4 秒, 且除了每个试次开始时的注视点外, 每两屏间都还有一个 600~800 ms 的注视点空屏, 试次结束后还有一个 800~1200 ms 的注视点空屏。学习阶段包括 100 试次, 测试阶段包括 200 试次。完成练习试次后, 先进行九点模式校准, 要求平均误差低于 0.5°, 成功校准后, 实验开始。

在学习阶段的价值反馈屏收集眼动数据。一方面, 以图片和分值区域为兴趣区, 分别收集被试的注视时间; 另一方面, 记录该屏内被试的瞳孔直径。

3.2 结果分析

3.2.1 奖赏预测误差对情景记忆成绩的影响

删去学习阶段被试没有做出价值猜测的试次后, 不同 RPE 条件下, 项目和联结记忆击中率的均值和标准差如表 6 所示, RPE 和奖赏结果对记忆击中率的影响如图 5 所示。

对记忆成绩进行分析(表 7)。对于项目记忆, 首先, 奖赏结果(p < 0.001)和 RPE (p < 0.001)主效应均显著, 且两者存在显著的负向交互作用(p = 0.002), 高奖赏结果和 RPE 负效价都促进了项目记忆的提高, 但效价效应会随着奖赏结果的增大而被削弱。其次, URPE 的影响也显著(p < 0.001), 项目记忆成绩随 RPE 凸显性增高而提高。联结记忆成绩的结果与实验 1 相同, 奖赏结果(p < 0.001)、RPE (p = 0.002)和 URPE (p < 0.001)的影响均显著, 且 RPE 与奖赏结果的负向交互作用显著(p < 0.001), 联结记忆成绩受到高奖赏结果的促进, 且存在 RPE 正效价和低凸显性优势, 其中效价效应会受高奖赏结果的削弱。

对再认正确试次的反应时进行分析(表 8), 只发现了 URPE 对联结记忆反应时的显著影响(p = 0.011), RPE 凸显性的减小促进了联结记忆提取速度的提高。

奖惩敏感性差值与正负效价间项目记忆成绩差值(r = −0.05, p = 0.840)和联结记忆成绩差值(r = −0.11, p = 0.630) 的相关都不显著。

3.2.2 奖赏预测误差对编码阶段眼动的影响

删去眼动数据收录过程中校准失败的 3 名被试, 对于剩余 20 名的被试, 删去眨眼过多或对分值注视时间小于 50 ms 的试次, 描述统计结果如表 9 所

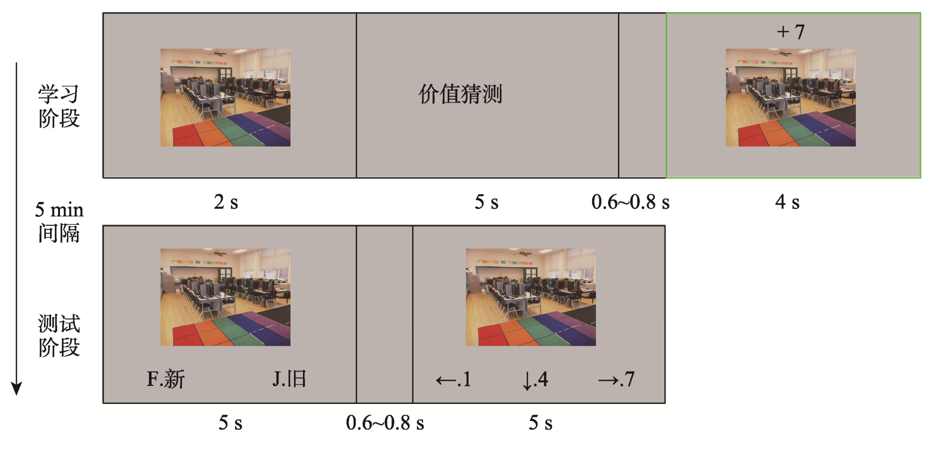


图 4　实验 2 流程图(在价值反馈屏收集眼动指标, 流程图中以绿色框标出)

表 6 实验 2 项目和联结记忆击中率均值和标准差($n = 23$)

记忆类型	RPE				
	−6	−3	0	3	6
项目记忆	0.85 (0.11)	0.75 (0.16)	0.74 (0.13)	0.78 (0.17)	0.81 (0.18)
联结记忆	0.33 (0.21)	0.31 (0.14)	0.51 (0.14)	0.38 (0.18)	0.43 (0.21)

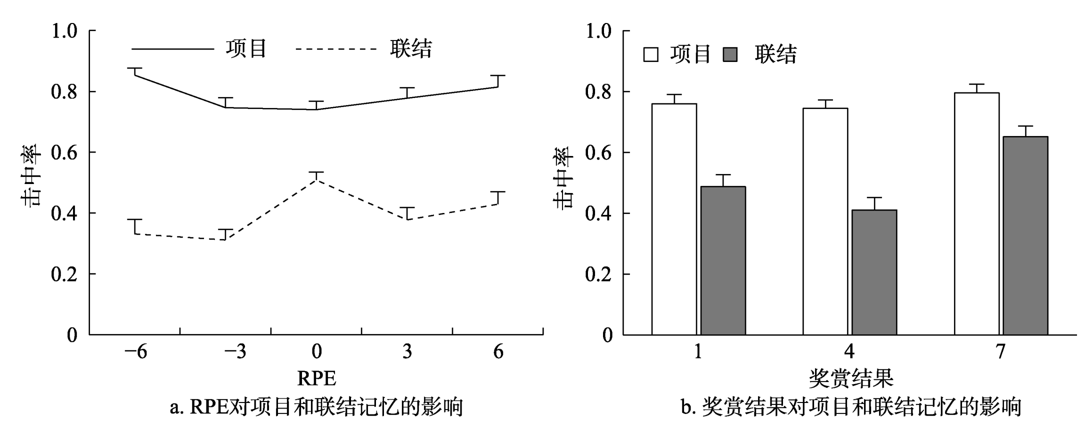


图 5 实验 2 奖赏(a 为 RPE, b 为奖赏结果)对项目和联结记忆击中率的影响(误差线为标准误)

表 7 实验 2 项目和联结记忆击中率广义混合线性模型分析结果

变量	项目记忆击中率				联结记忆击中率			
	b	*z*	*p*	95% CI	*b*	*z*	*p*	95% CI
intercept	0.82	4.09	<0.001	[0.53, 1.06]	−0.14	−1.03	0.302	[−0.42, 0.13]
RPE	−0.15	−4.16	<0.001	[−0.22, −0.08]	0.10	3.07	0.002	[0.04, 0.17]
奖赏结果	0.09	3.38	<0.001	[0.04, 0.15]	0.12	4.25	<0.001	[0.06, 0.17]
交互作用	0.02	3.04	0.002	[0.01, 0.04]	−0.03	−4.06	<0.001	[−0.04, −0.01]
intercept	0.96	4.99	<0.001	[0.58, 1.35]	0.20	1.60	0.110	[−0.05, 0.44]
URPE	0.09	3.57	<0.001	[0.04, 0.13]	−0.18	−7.72	<0.001	[−0.22, −0.13]
奖赏结果	0.04	2.11	0.035	[0.00, 0.08]	0.10	5.21	<0.001	[0.06, 0.14]

表 8 实验 2 项目和联结记忆反应时广义混合线性模型分析结果

变量	项目记忆反应时(ms)				联结记忆反应时(ms)			
	b	*t*	*p*	95% CI	*b*	*t*	*p*	95% CI
intercept	1643.47	24.73	<0.001	[1513.13, 1773.81]	898.77	11.10	<0.001	[739.81, 1057.73]
RPE	−1.85	−0.31	0.760	[−13.73, 10.03]	−4.04	−0.48	0.634	[−20.70, 12.62]
奖赏结果	−12.14	−1.500	0.134	[−28.02, 3.73]	−13.23	−1.30	0.195	[−33.26, 6.80]
intercept	1656.91	25.55	<0.001	[1529.73, 1784.09]	854.83	10.84	<0.001	[700.02, 1009.64]
URPE	−1.98	−0.299	0.765	[−14.95, 10.99]	23.01	2.55	0.011	[5.27, 40.76]
奖赏结果	−14.08	−2.423	0.016	[−25.48, −2.68]	−14.63	−1.86	0.063	[−30.08, 0.82]

表 9 实验 2 注视时间(ms)和瞳孔变化(μm)的均值和标准差($n = 20$)

眼动指标	RPE				
	−6	−3	0	3	6
图片注视时间	2660.54 (460.5)	2576.9 (403.95)	2572.34 (356.85)	2686.05 (367.19)	2721.83 (373.85)
分值注视时间	506.59 (224.34)	542.96 (212.07)	550.89 (246.87)	486.1 (202.6)	437.07 (176.77)
瞳孔变化平均值	−69.78 (67.04)	−42.83 (62.85)	−33.84 (48.4)	−44.19 (69.11)	−42.25 (82.59)
瞳孔变化最大值	159.72 (57.69)	174.92 (61.73)	176.47 (57.2)	180.08 (58.26)	210.06 (98.49)

表 10　实验 2 眼动结果的混合线性模型分析结果

变量	图片注视时间(ms)				分值注视时间(ms)			
	b	t	p	95% CI	b	t	p	95% CI
intercept	2607.41	30.35	<0.001	[2438.91, 2775.92]	528.31	10.53	<0.001	[429.88, 626.74]
RPE	8.72	1.46	0.143	[−2.96, 20.41]	−5.72	−1.50	0.133	[−13.16, 1.74]
奖赏结果	1.98	0.25	0.801	[−13.43, 17.39]	−2.09	−0.42	0.676	[−11.92, 7.73]
intercept	2531.79	29.81	<0.001	[2365.20, 2698.39]	851.26	11.77	<0.001	[484.36, 678.15]
URPE	15.66	2.41	0.016	[2.93, 28.39]	−11.51	−2.78	0.006	[−19.62, −3.39]
奖赏结果	10.81	1.90	0.058	[−0.36, 21.97]	−7.95	−2.19	0.029	[−15.06, −0.83]

变量	瞳孔变化平均值(μm)				瞳孔变化最大值(μm)			
	b	t	p	95% CI	b	t	p	95% CI
intercept	−25.94	−1.84	0.067	[−53.66, 1.79]	192.32	14.50	<0.001	[166.31, 218.33]
RPE	7.06	3.05	0.002	[2.53, 11.60]	4.40	3.63	<0.001	[2.03, 6.78]
奖赏结果	−2.08	−1.11	0.266	[−5.74, 1.58]	−3.33	−2.08	0.037	[−6.47, −020]
交互作用	−1.09	−2.28	0.023	[−2.02, −0.15]				
intercept	−21.52	−1.46	0.144	[−50.40, 7.37]	194.58	13.95	<0.001	[167.22, 221.95]
URPE	−7.84	−2.97	0.003	[−13.02, −2.66]	−6.67	−2.95	0.003	[−11.11, −2.23]
奖赏结果	−2.80	−1.38	0.169	[−6.79, 1.19]	−4.28	−2.45	0.014	[−7.70, −0.86]
交互作用	1.18	2.12	0.034	[0.09, 2.28]	1.83	3.82	<0.001	[0.89, 2.77]

示。对图片和分值注视时间以及瞳孔直径变化的混合线性模型分析结果如下(见表 10)。

考察 RPE 对图片和分值注视时间的影响。结果发现，URPE 对图片注视时间影响显著(p = 0.016)，RPE 凸显性的增大会使图片注视时间延长。URPE 对分值注视时间也有显著影响(p = 0.006)，RPE 凸显性更小时，被试对分值的注视时间越长。奖赏结果和 RPE 对图片注视时间的影响都不显著，奖赏结果和 RPE 对分值注视时间的影响也都不显著(ps > 0.010)。

首先进行瞳孔基线校正，选取价值反馈屏呈现前 500 ms 的平均瞳孔直径作为基线，通过减法换算(瞳孔变化值 = 反馈屏瞳孔大小 − 基线值)，得到校正后的瞳孔变化平均值和最大值(杨晓梦 等, 2020)。考察 RPE 对瞳孔变化的影响，结果发现，对于瞳孔变化平均值，RPE (p = 0.002)和 URPE (p = 0.003)的主效应显著，且奖赏结果和 RPE 有显著正向交互作用(p = 0.023)，奖赏结果和 URPE 也有显著正向交互作用(p = 0.034)，表明 RPE 正效价和低凸显性都促进了瞳孔变化平均值的增加，同时奖赏结果的增大会增强效价和凸显性的效应。对于瞳孔变化最大值，奖赏结果(p = 0.037)、RPE (p < 0.001)和 URPE (p = 0.003)的主效应显著，且奖赏结果和 URPE 有显著正向交互作用(p < 0.001)，表明 RPE 正效价和低凸显性都促进了瞳孔变化最大值的增加，同时奖赏结果的增大会增强凸显性的效应。

3.3　讨论

在缩短预期阶段时长，并延长反馈阶段的呈现时间之后，实验 2 发现，项目记忆中存在 RPE 负效价与高凸显性优势，与我们的预期相符。项目记忆对价值反馈阶段认知资源的竞争并没有改变联结记忆中的 RPE 效应，实验 2 再次验证了联结记忆中的 RPE 正效价和低凸显性优势。

眼动指标反映了元记忆控制过程，其中注视时间反映了个体对不同刺激的学习时间分配(姜英杰 等, 2016)，主要受到了 RPE 凸显性的影响。当 RPE 凸显性更低时，被试采取优先编码联结信息的策略，延长了对分值的学习时间，此时对应的联结记忆成绩也更高。而当 RPE 凸显性增大，奖赏联结的编码优先级降低，被试对图片的学习时间相应延长。

瞳孔直径变化体现了被试的心理努力程度(Ariel & Castel, 2014)。校正后的瞳孔变化平均值和最大值结果都表明，RPE 正效价和低凸显性促进了瞳孔变化，同时奖赏结果更高时，RPE 对瞳孔变化的影响会更大。这与 RPE 对联结记忆的影响基本相符，可能表明瞳孔直径的变化主要反映了个体在奖赏动机作用(参见综述：杨晓梦 等, 2020)下对分值加工的心理努力程度。

实验 1 和 2 发现了 RPE 效价和凸显性对项目和联结记忆的分离影响，并进一步揭示了该影响发生的认知机制——奖赏信息通过元记忆控制改变了

个体对项目和联结编码的认知资源投入。但在这两个实验中，奖赏结果只包含 3 个水平，导致高低奖赏结果分别与正负效价 RPE 重叠较大，虽然统计分析中始终考虑了奖赏结果的影响，已经较准确地体现了 RPE 效价本身的效应，但为增强该结果的信度，补充进行实验 3，目的是考察实验 1 和 2 中 RPE 效价对记忆成绩影响的稳定性，并再次验证 RPE 凸显性的影响。

4 实验 3：奖赏预测误差对项目和联结记忆成绩影响的稳定性

实验 1 和 2 中，奖赏结果只有 1、4 和 7 三种，导致高奖赏结果 7 对应的 RPE 只有 0 和负效价，低奖赏结果 1 对应的 RPE 只有 0 和正效价，只有中等奖赏结果 4 同时对应正效价、0 和负效价。因此在实验 3 中，将奖赏结果设置为 1、3、5 和 7 四种，以增加 RPE 水平，减小 RPE 效价和高低奖赏结果的重叠，验证 RPE 效价和凸显性对项目和联结记忆影响的稳定性。

4.1 方法

4.1.1 被试

招募被试 29 人，剔除 2 名项目记忆击中率低于 0.5 的被试，最后保留被试 27 人。其中女性 24 人，年龄 19~25 岁(M = 22.11 岁，SD = 1.98 岁)，视力或矫正视力正常，无神经系统和心理疾病病史。所有被试均自愿参加实验，并签署被试知情同意书。每名被试获得真实现金奖励平均约为 19 元。

4.1.2 实验材料

刺激材料和量表与前两个实验相同。

4.1.3 实验设计与流程

实验设计与实验 1 相似。但在自变量的操纵上，奖赏结果序列变为 1、3、5 和 7，它们出现的比例在高价值图片中是 1 : 2 : 3 : 4，而在低价值图片中为 4: 3: 2: 1。由奖赏结果和被试的给出的猜测价值之差产生 RPE 共有−6、−4、−2、0、2、4 和 6 七种。

总体实验流程和指导语与实验 2 相似。不同的是，首先，只保留了试次开始时的注视点屏；其次，考虑到分值序列增加对记忆难度的改变，将学习阶段价值反馈屏的呈现时间延长到 5 秒。学习阶段包括 100 试次，测试阶段包括 200 试次。

4.2 结果分析

删去学习阶段被试没有做出价值猜测按键的试次后，不同 RPE 条件下，项目和联结记忆击中率的均值和标准差如表 11 所示，RPE 和奖赏结果对记忆击中率的影响如图 6 所示。

对记忆成绩进行分析(表 12)。对于项目记忆，与实验 2 相同，奖赏结果(p = 0.001)、RPE (p < 0.001)和 URPE (p < 0.001)的影响均显著，且 RPE 和奖赏结果有显著的负向交互作用(p < 0.001)，奖赏结果的增大促进了项目记忆成绩的提高，项目记忆存在 RPE 负效价优势以及高凸显性优势，但奖赏结果的提高会削弱效价效应。联结记忆的结果与前两个实

表 11 实验 3 项目和联结记忆击中率均值和标准差(n = 27)

记忆类型	RPE						
	−6	−4	−2	0	2	4	6
项目记忆	0.85(0.19)	0.83(0.14)	0.81(0.12)	0.76(0.16)	0.76(0.16)	0.83(0.16)	0.89(0.23)
联结记忆	0.18(0.27)	0.26(0.19)	0.34(0.12)	0.46(0.11)	0.33(0.12)	0.28(0.21)	0.26(0.29)

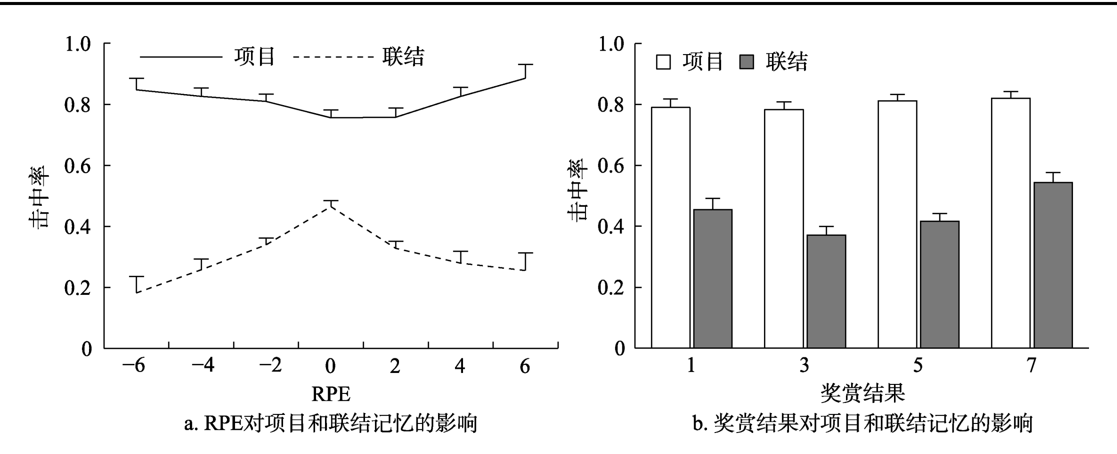


图 6 实验 3 奖赏(a 为 RPE, b 为奖赏结果)对项目和联结记忆击中率的影响(误差线为标准误)

表 12　实验 3 项目和联结记忆击中率广义混合线性模型分析结果

变量	项目记忆击中率				联结记忆击中率			
	b	*z*	*p*	95% CI	*b*	*z*	*p*	95% CI
intercept	1.02	5.61	<0.001	[0.66, 1.39]	−0.18	−1.48	0.14	[−0.41, 0.06]
RPE	−0.21	−4.48	<0.001	[−0.30, −0.12]	0.21	5.14	<0.001	[0.13, 0.29]
奖赏结果	0.09	3.23	0.001	[0.04, 0.15]	0.05	1.91	0.056	[−0.00. 0.10]
交互作用	0.04	3.64	<0.001	[0.02, 0.06]	−0.05	−6.00	<0.001	[−0.07, −0.03]
intercept	1.11	6.31	<0.001	[0.76, 1.46]	0.18	1.67	0.095	[−0.03, 0.40]
URPE	0.12	3.81	<0.001	[0.06, 0.17]	−0.28	−10.23	<0.001	[−0.33, −0.23]
奖赏结果	0.05	2.14	0.032	[0.00, 0.09]	0.05	2.51	0.012	[0.01, 0.09]

表 13　实验 3 项目和联结记忆反应时混合线性模型分析结果

变量	项目记忆反应时(ms)				联结记忆反应时(ms)			
	b	*t*	*p*	95% CI	*b*	*t*	*P*	95% CI
intercept	1704.79	26.76	<0.001	[1579.85, 1829.74]	1156.36	12.21	<0.001	[970.49, 1342.22]
RPE	−8.22	−1.17	0.244	[−22.03, 5.60]	−43.81	−3.26	0.001	[−70.18, −17.44]
奖赏结果	5.77	0.68	0.500	[−10.83, 22.37]	8.18	0.62	0.533	[−17.54, 33.91]
intercept	1712.91	27.53	<0.001	[1590.88, 1834.95]	1052.67	10.37	<0.001	[853.47, 1251.86]
URPE	8.42	1.02	0.308	[−7.77, 24.62]	121.64	4.16	<0.001	[64.33, 178.95]
奖赏结果	−0.30	−0.05	0.963	[−12.85, 12.25]	4.99	0.34	0.737	[−24.15, 34.13]
交互作用					−11.83	−1.97	0.049	[−23.58, −0.07]

验相似, RPE ($p < 0.001$)和 URPE ($p < 0.001$)的影响均显著, 且存在 RPE 与奖赏结果显著的负向交互作用($p < 0.001$), 但奖赏结果的影响只达到了边缘显著水平($p = 0.056$), 表明联结记忆成绩存在稳定的 RPE 正效价和低凸显性优势, 存在随奖赏结果增大而提高的趋势, 且其中的 RPE 效价效应还会受到高奖赏结果的削弱。

对再认正确试次的反应时进行分析(表 13), RPE ($p = 0.001$)和 URPE ($p < 0.001$)对联结记忆反应时的影响均显著, 且奖赏结果与 URPE 的交互作用显著($p = 0.049$), 表明联结记忆的提取速度在 RPE 正效价或凸显性更低时加快, 且奖赏结果增大后会削弱凸显性的效应。

奖惩敏感性差值与正负效价间项目记忆成绩差值($r = 0.02$, $p = 0.940$)和联结记忆成绩差值($r = 0.18$, $p = 0.370$)的相关都不显著。

4.3　讨论

将 RPE 的水平数增加到 7 后, 正负效价与高低奖赏结果的重叠减小, 但 RPE 对项目和联结记忆成绩的影响与实验 1 和 2 中相同, 这表明 RPE 效价和凸显性对项目和联结记忆的影响是稳定存在的, RPE 效价的影响并非是由奖赏结果差异引起的假阳性结果。

对于联结记忆提取过程, 实验 1 中发现了高奖赏结果对提取流畅性的促进, 而实验 3 将 RPE 效价与奖赏结果高低的影响进一步分离后, 发现 RPE 正效价时联结记忆提取速度更快, 同时奖赏结果的影响不再显著。这表明实验 1 中的奖赏结果的影响可能源自于 RPE 的效价效应。

实验 2 和 3 中均未发现奖惩敏感性与项目和联结记忆中 RPE 效价效应的相关, 且实验 1 中的相关程度偏低, 这可能是由于受招募而来的被试群体对奖惩的敏感性类似。通过将标准差除以平均数计算离散系数, 发现被试对奖赏和惩罚的敏感性原始分数的离散系数均低于 0.5 (实验 1, 奖赏：0.43, 惩罚：0.40; 实验 2, 奖赏：0.46, 惩罚：0.29; 实验 3, 奖赏：0.15, 惩罚：0.17), 离散程度处于较低水平, 验证了上述推论。

5　总讨论

基于价值学习−测试范式, 本研究通过 3 个实验探讨了 RPE 效价和凸显性对同一任务中的项目和联结记忆的不同影响。实验 1 发现, 联结记忆成绩存在 RPE 正效价和低凸显性优势, 其 JOCs 准确性在 RPE 正效价时更高。实验 2 进一步发现项目记忆成绩存在与联结记忆相反的 RPE 负效价和高

凸显性优势，且通过眼动技术考察记忆编码中的元认知控制过程发现，RPE 正效价和低凸显性提高了个体的瞳孔变化平均值和峰值，且 RPE 低凸显性增加了分值注视时间，缩短了图片注视时间。实验 3 增加了 RPE 水平，再次验证了 RPE 对项目和联结记忆的相反影响。

5.1 奖赏预测误差对项目和联结记忆成绩影响的分离

RPE 效价对项目和联结记忆存在分离的影响。项目记忆成绩中的 RPE 负效价优势与 Wimmer 等(2014)的发现一致。RPE 正效价能够促进联结记忆成绩的提高，与使用词对联结(Ergo et al., 2020)、名字–面孔联结(Calderon et al., 2021)和面孔–物品图像联结(Aberg et al., 2017)的前人实验结果一致，体现了联结记忆中 RPE 正效价优势的类别一般性。

RPE 凸显性对项目和联结记忆也存在分离的影响。RPE 高凸显性对项目记忆的促进作用与前人相同(Rouhani & Niv, 2019, 2021; Rouhani et al., 2018)。RPE 高凸显性对联结记忆的抑制作用也与 Rouhani 等(2020)的结果相符。而且与前人研究不同，本研究中的项目–奖赏联结记忆涉及实验的主要任务，受到奖赏的直接影响，联结记忆中的凸显性效应体现的是 RPE 凸显性的直接影响。此外，在实验 1 中，项目记忆在不同奖赏条件下的成绩相似，占用了相似的认知资源，而联结记忆的高凸显性优势仍然存在，再次表明 RPE 凸显性是直接作用于联结记忆，而非通过影响不同凸显性时项目记忆对认知资源的占用间接产生。

5.2 奖赏预测误差影响记忆编码中的元记忆控制

5.2.1 效价对元记忆控制的影响

情景记忆中的奖赏联结记忆(Murty et al., 2016)能够引导适应性决策，在认知层面，当面对意味着更高收益的 RPE 正效价时，个体的瞳孔变化增大，这可能体现了个体有了更强的奖赏动机后，主动通过元记忆控制增强对联结记忆的认知资源分配的过程。根据注意的双重竞争模型(Pessoa, 2009)，一方面，奖赏相关刺激在视觉皮层上的表征得到增强，能够被优先注意；另一方面，个体会通过增强执行功能，促进奖赏相关加工的效率。此外，在生理层面，RPE 正效价促使腹侧被盖区的多巴胺分泌增加，而负效价时多巴胺分泌减少(Schultz et al., 1997)，这些多巴胺传入海马后能够使瞬时记忆痕迹获得更大的时间持久性，从而在联结记忆编码时或编码前后产生促进作用(Bethus et al., 2010)。

而 RPE 负效价时，对联结记忆的资源投入减少，导致此时联结记忆成绩降低。但对于项目记忆，有核磁研究发现，在编码阶段，RPE 负效价虽然对纹状体的激活更弱，但负效价的图片的再认正确率比正效价图片更高；且再认成功的图片在编码中对应的双侧海马激活更大(Wimmer et al., 2014)。这表明，RPE 负效价虽然诱发了更弱的奖赏动机，却会促进项目记忆编码，从而提高了项目记忆表现。

5.2.2 凸显性对元记忆控制的影响

眼动结果表明，凸显性对图片和分值编码时长的影响是“此消彼长”的，这一元记忆控制过程体现了项目和奖赏联结对认知资源的竞争，也是二者受 RPE 凸显性的影响表现出分离形式的主要原因。RPE 凸显性代表结果与预期的偏离程度的大小，低凸显性意味着更少的分值信息更新负荷，降低了联结编码难度，加工流畅性更高。而加工流畅性作为编码过程元记忆监测的内在线索(Koriat, 1997)，会促使被试采取优先对联结信息分配更多认知资源的元记忆控制策略。一方面，延长对分值的学习时间，以加深奖赏联结的记忆痕迹，另一方面，此时瞳孔变化增强，表明了被试增加了联结编码的心理努力程度。

RPE 凸显性增大后，联结记忆加工流畅性降低，奖赏联结的优先级降低，对图片的认知资源分配相应增加。在生理层面，此时被试的意外程度更大，可能会引起蓝斑中去甲肾上腺素分泌的增加，去甲肾上腺素与多巴胺在海马中共同释放能够调节编码后记忆增强(Takeuchi et al., 2016)，促进了此时的项目记忆编码。

5.3 奖赏预测误差影响联结提取中的元记忆监测

联结记忆提取过程的元记忆监测准确性受到了 RPE 正效价的促进，流畅性受到 RPE 正效价和低凸显性的促进。根据记忆再认的双加工理论，联结记忆提取需要检索特定项目的情景信息，只能通过更复杂的回想过程完成，不能像项目记忆一样通过熟悉性过程而快速地、自动化地完成再认(Woroch & Gonsalves, 2010)。因此编码阶段中的奖赏在提取阶段的效应主要体现在对回想过程的促进。

研究者发现积极情绪刺激在编码阶段会被赋予更多认知资源，促进对视觉细节的加工，从而促进情景记忆的回想过程(毛新瑞 等, 2015)。而相比于基于熟悉性做出信心判断，当个体的信心判断是基于回想时，元记忆监测更准确(Souchay et al., 2013)。由于比预期更高的奖赏结果通常会引起积

极情绪, 这可能解释了奖赏对联结记忆提取中的元记忆监测产生促进作用的机制——RPE 正效价通过在编码阶段中促进对细节的加工, 增强了联结提取中的回想过程, 从而提高了元记忆监测准确性。

5.4　高奖赏结果同时促进项目和联结记忆

越来越多的研究者认为奖赏结果主要是通过 RPE 产生作用, 因而着重探索 RPE 对记忆的影响, 同时忽略了奖赏结果的作用(Ergo et al., 2020; Jang et al., 2019; Rouhani et al., 2018)。但本研究结果表明, 奖赏结果在项目和联结记忆中都发挥着重要作用, 尤其当结果令人足够满意时, 结果与预期孰优孰劣将不再作为个体判断其损益的主要标准。

综合三个实验结果, 奖赏结果的增大同时促进了项目记忆和联结记忆的提高。奖赏结果能够正向促进情景记忆成绩, 这与前人研究的结果符合(Castanheira et al., 2021)。但 Murty 等(2016)发现面孔−低奖赏联结的记忆成绩更好, 这可能是因为他们使用具有社会性的独裁者游戏范式, 被试采取“避害”的记忆策略。总的来说, 奖赏结果对记忆的影响具有“趋利避害”的适应性。

5.5　不足与展望

本研究考察了 RPE 效价和凸显性如何同时作用于项目和联结记忆, 发现二者在编码阶段中存在认知资源竞争, 被试会以 RPE 效价和凸显性为线索对其进行适应性的选择性加工。在课堂的记忆促进中, 对奖赏的利用通常局限于奖赏结果本身, 但本研究结果表明, 若同时操纵了学习者的奖赏预期, 使其形成不同的 RPE, 对记忆效果可能有更多的促进作用。

但值得注意的是, 本研究选取的联结记忆指标为奖赏联结, 与前人研究中的人物−物品联结(Aberg et al., 2017)以及外语−母语词汇联结(Calderon et al., 2021; Ergo et al., 2021)等存在一个共同点, 即 RPE 凸显性大小对应了不同的联结编码难度, 这可能是导致 RPE 低凸显性时联结记忆成绩升高的重要原因。因此, 当联结编码的难度与凸显性无关时, 凸显性的影响是否仍存在, 是未来研究需要考虑的问题。

个体本身的特质能够影响奖赏的动机作用从而作用于项目记忆(Rouhani & Niv, 2019)和联结记忆(Aberg et al., 2017), 但本研究只在实验 1 中发现了奖惩敏感性与联结记忆中 RPE 效价效应的低相关, 这可能是由被试奖惩敏感性的低离散程度导致的。之后的研究可以考虑在实验前选取高低敏感性两组被试, 探索其情景记忆受奖赏动机影响的差异。

此外, 实验 2 通过眼动技术对记忆编码过程的学习时间分配和瞳孔变化表征的心理努力程度进行了考察, 发现了编码项目和联结信息的认知资源分配随 RPE 而变化。但因为瞳孔直径容易受物理环境中的光刺激影响, 而 EEG 研究能在时间维度更精细地反映记忆编码过程中的神经激活程度, 未来的研究可以通过分析 RPE 影响情景记忆的 EEG 时程变化, 进一步验证两者的编码过程受奖赏信息影响的差异。

6　结论

RPE 对项目和联结记忆存在分离的影响, 元记忆过程在其中发挥着重要作用。在编码阶段, 个体以 RPE 为线索进行元记忆控制, RPE 高凸显性增加了对项目编码的认知资源投入, RPE 正效价和低凸显性增加了对联结编码的认知资源投入; 在提取阶段, RPE 正效价还通过回想过程促进联结记忆提取的元记忆监测准确性。

参　考　文　献

Aberg, K. C., Muller, J., & Schwartz, S. (2017). Trial-by-trial modulation of associative memory formation by reward prediction error and reward anticipation as revealed by a biologically plausible computational model. *Frontiers in Human Neuroscience*, *11*, 56.

Ariel, R., & Castel, A. D. (2014). Eyes wide open: Enhanced pupil dilation when selectively studying important information. *Experimental brain research*, *232*(1), 337−344.

Bethus, I., Tse, D., & Morris, R. G. M. (2010). Dopamine and memory: Modulation of the persistence of memory for novel hippocampal nmda receptor-dependent paired associates. *Journal of Neuroscience*, *30*(5), 1610−1618.

Calderon, C. B., De Loof, E., Ergo, K., Snoeck, A., Boehler, C. N., & Verguts, T. (2021). Signed reward prediction errors in the ventral striatum drive episodic memory. *Journal of Neuroscience*, *41*(8), 1716−1726.

Da Silva Castanheira, K., Lalla, A., Ocampo, K., Otto, A. R., & Sheldon, S. (2021). Reward at encoding but not retrieval modulates memory for detailed events. *Cognition*, *219*, 104957.

Ergo, K., De Loof, E., Debra, G., Pastotter, B., & Verguts, T. (2020). Failure to modulate reward prediction errors in declarative learning with theta (6 Hz) frequency transcranial alternating current stimulation. *Plos One*, *15*(12), e0237829.

Ergo, K., De Vilder, L., De Loof, E., & Verguts, T. (2021). Reward prediction errors drive declarative learning irrespective of agency. *Psychonomic Bulletin & Review*, *28*(6), 2045−2056.

Guo, Y., Song, G., Zhao, P., & Ma, Y. (2011). Revision of the sensitivity to punishment and sensitivity to reward questionnaire. *Journal of Jinan Vocational College*, (1), 91−94+97.

[郭永香, 宋广文, 赵平平, 马玉花. (2011). 大学生惩罚和奖励敏感性问卷(SPSRQ)的修订. *济南职业学院学报*, (1),

91−94+97.]

Höltje, G., & Mecklinger, A. (2020). Feedback timing modulates interactions between feedback processing and memory encoding: Evidence from event-related potentials. *Cognitive Affective & Behavioral Neuroscience*, *20*(2), 250−264.

Jang, A. I., Nassar, M. R., Dillon, D. G., & Frank, M. J. (2019). Positive reward prediction errors during decision-making strengthen memory encoding. *Nature Human Behaviour*, *3*(7), 719−732.

Jiang, Y., Wang, Z., Zheng, M., & Jin, X. (2016). How value-based agendas affect study time allocation: An eye tracking study. *Acta Psychologica Sinica*, *48*(10), 1229−1238.

[姜英杰, 王志伟, 郑明玲, 金雪莲. (2016). 基于价值的议程对学习时间分配影响的眼动研究. *心理学报*, *48*(10), 1229−1238.]

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349−370.

Mao, X., Xu, H., & Guo, C. (2015). Emotional memory enhancement effect in dual-processing recognition retrieval. *Acta Psychologica Sinica*. *47*(9), 1111−1123.

[毛新瑞, 徐慧芳, 郭春彦. (2015). 双加工再认提取中的情绪记忆增强效应. *心理学报*, *47*(9), 1111−1123.]

Mason, A., Lorimer, A., & Farrell, S. (2019). Expected value of reward predicts episodic memory for incidentally learnt reward-item associations. *Collabra-Psychology*, *5*(1), 40.

Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*, *145*(5), 548−558.

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160−166.

Rouhani, N., & Niv, Y. (2019). Depressive symptoms bias the prediction-error enhancement of memory towards negative events in reinforcement learning. *Psychopharmacology*, *236*(8), 2425−2435.

Rouhani, N., & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory. *Elife*, *10*, e61077.

Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, *44*(9), 1430−1443.

Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020). Reward prediction errors create event boundaries in memory. *Cognition*, *203*, 104269.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593−1599.

Souchay, C., Guillery-Girard, B., Pauly-Takacs, K., Wojcik, D. Z., & Eustache, F. (2013). Subjective experience of episodic memory and metacognition: A neurodevelopmental approach. *Front Behav Neurosci*, *7*, 212.

Takeuchi, T., Duszkiewicz, A. J., Sonneborn, A., Spooner, P. A., Yamasaki, M., Watanabe, M.,…., Morris, R. G. M. (2016). Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature*, *537*(7620), 357-+.

Tulving, E. (2001). Episodic memory and common sense: How far apart? *Philosophical Transactions of the Royal Society B-Biological Sciences*, *356*(1413), 1505−1515.

Wimmer, G. E., Braun, E. K., Daw, N. D., & Shohamy, D. (2014). Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. *Journal of Neuroscience*, *34*(45), 14901−14912.

Woroch, B., & Gonsalves, B. D. (2010). Event-related potential correlates of item and source memory strength. *Brain Research*, *1317*, 180−191.

Yang, X., Wang F., Wang, Y., Zhao, T., Gao, C., & Hu, X. (2020). Are pupils the window of our mind? Pupil-related application in psychology and pupillometry. *Advances in Psychological Science*, *28*(7), 1029−1041.

[杨晓梦, 王福兴, 王燕青, 赵婷婷, 高春颖, 胡祥恩. (2020). 瞳孔是心灵的窗口吗?——瞳孔在心理学研究中的应用及测量. *心理科学进展*, *28*(7), 1029−1041.]

The distinct effects of reward prediction error on item and associative memory: The influence of metamemory

LONG Yiting[1], JIANG Yingjie[1], CUI Can[2], YUE Yang[1]

([1] *School of Psychology, Northeast Normal University, Changchun* 130024, *China*)

([2] *Jiangsu Key Laboratory of Brain Disease and Bioinformation, Research Center for Biochemistry and Molecular Biology, Xuzhou Medical University, Xuzhou* 221004, *China*)

Abstract

Episodic memory consists of item memory and associative memory. Individual cognitive resources are typically allocated to more valuable information during encoding through metamemory, leading to competitive processing of item and associative information. Reward prediction error (RPE), defined as the difference between reward results and reward expectations, has two properties: valence (positive or negative) and salience (degree of difference). To examine the impact of reward prediction error valence and salience on item and associative memory, and how reward prediction error influences memory based on metamemory, three experiments were conducted.

In the learning stage, participants were presented with indoor and outdoor scene pictures. They were asked

to predict the score of each picture and then received feedback on the actual score. Through this reinforcement learning process, participants had to find out which type of pictures is more valuable, and 30% of the scores were accumulated into the total score. To induce the effect of reward motivation on memory, participants were introduced to the opportunity to choose between two pictures and receive the value of the selected picture, although the actual program did not include a decision-making stage. After the learning stage, participants were tested on item and reward associative memory.

The findings of the study showed that: (1) There were advantages in associative memory performance for positive reward prediction error valence and low salience, with higher accuracy of JOCs at positive valence. In contrast, there were advantages in item memory performance for negative valence and high salience. (2) In the eye-tracking results during the encoding process, positive valence and low salience of reward prediction error resulted in increased mean and peak pupil dilation after feedback presentation, as well as longer value fixation duration and shorter picture fixation duration at low salience. (3) When the reward prediction error level was increased to reduce overlap between reward results and reward prediction error effects, the separation effect of reward prediction error on item and associative memory performance remained stable.

The results of the study suggest that the effects of reward prediction error on item and associative memory are distinct. During the encoding stage, individuals use the valence and salience of reward prediction error as cues to allocate cognitive resources differently in item and associative memory encoding through metamemory control. In the retrieval stage, positive valence of reward prediction error enhances the metamemory monitoring level of associative memory retrieval.

Keywords reward prediction error, associative memory, eye movements, episodic memory, metamemory